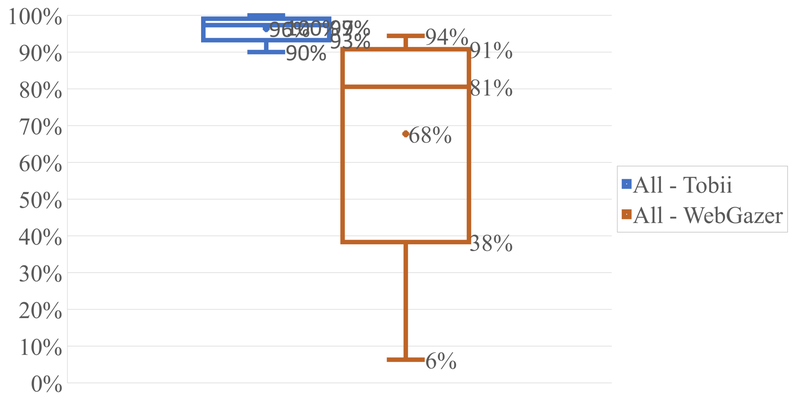
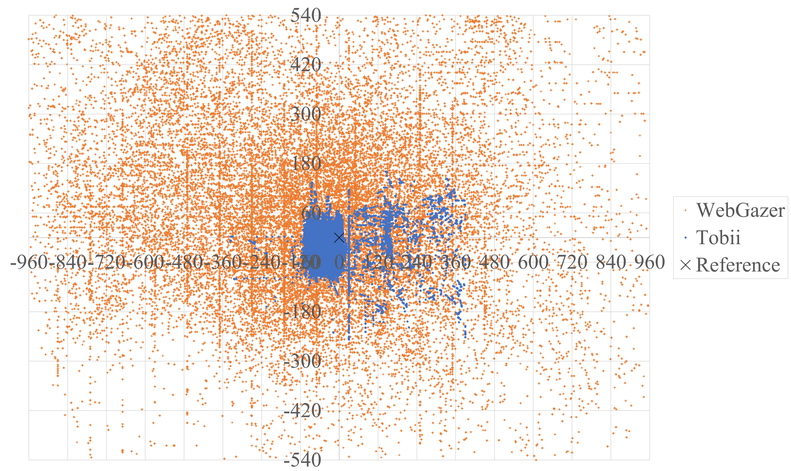
Comparative Analysis of RGB-based Eye-Tracking for Large-Scale Human-Machine Applications

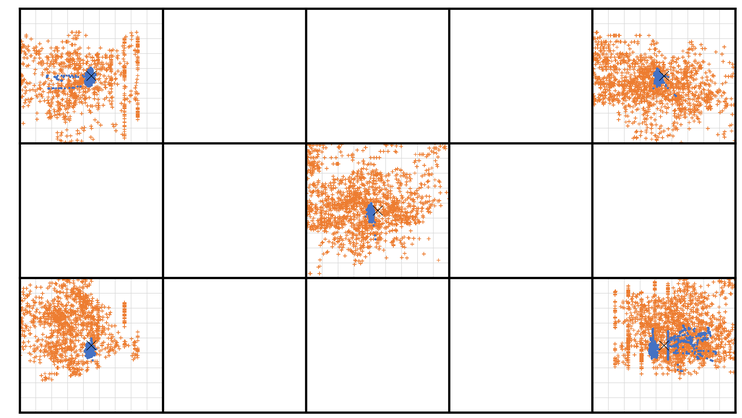
Brett Thaman[1], Trung Cao[1], Nicholas Caporusso[1] - {thamanb1, caot1, caporusson1}@nku.edu
[1] Department of Computer Science, Northern Kentucky University

Abstract
Gaze tracking has become an established technology that enables using an individual's gaze as an input signal to support a variety of applications in the context of Human-Computer Interaction. Gaze tracking primarily relies on sensing devices such as infrared (IR) cameras. Nevertheless, in the recent years, several attempts have been realized at detecting gaze by acquiring and processing images acquired from standard RGB cameras. Nowadays, there are only a few publicly available open-source libraries and they have not been tested extensively.

Objective and Methods
In this poster, we present the result of a comparative analysis that studied a commercial eye-tracking device using IR sensors, that is, Tobii 4C and WebGazer, a software system that uses machine learning and linear regression to estimate gaze from images acquired by a standard webcam.
A total of 14 people volunteered to participate in the experiment. We collected a total of 228336 data points representing: the coordinates of the reference (i.e., the position of the circle on the display), the gaze location acquired by Tobii, and the gaze location estimated by WebGazer.
Then, we calculated the distance in pixels between the position of the reference and the gaze predictions (i.e., dispersion). Subsequently, we calculated the overall performance in terms of accuracy, by normalizing the dispersion with respect to the size of the display.

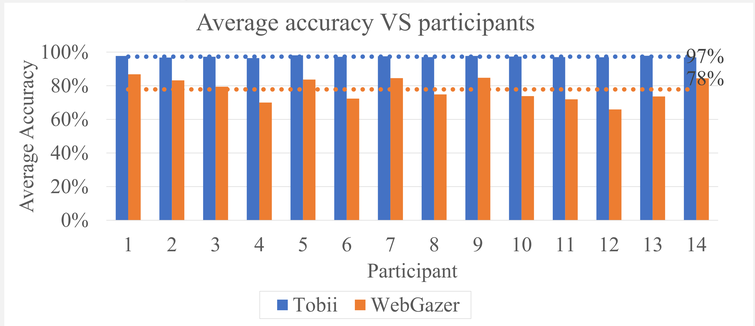
Accuracy


Dispersion – all data


Dispersion by screen quadrants


Dispersion by screen quadrants

Quadrant	Tobii	WebGazer
Top-left (q 1,1)	28.64±16.83 px	671.31±430.33 px
Bottom-left (q 3,1)	41.92±21.38 px	549.18±323.42 px
Center (q 2,3)	59.14±12.74 px	434.04±204.82 px
Top-right (q 1,5)	94.66±31.52 px	549.15±346.03 px
Bottom-right (q 3,5)	89.69±27.18 px	532.65±360.58 px

Cross-subject reliability


Conclusion
From our findings, we can conclude that, despite the advancements in artificial intelligence and computer vision, gaze tracking using IR sensors is still significantly more accurate than RGB webcams. Specifically, the software library tested in our work is not suitable for gaze tracking tasks that require accuracy and reliability.

Acknowledgements
The project received support from NKU's Institute for Health Innovation (IHI) through the IHI Faculty Fellows initiative.

Project repository
https://bit.ly/3DBiQFc